TDT4173 - Machine Learning and case-based reasoning

Assignment One

Joachim Krüger [joachibe@stud.ntnu.no]

September 5, 2012

Contents

1 Problem One

T. Mitchell defines a well posed learning problem as

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improve by experience E.

1.1

In mathematics a well posed problem is one that has a (unique) solution, and where a small change in the problem's initial condition(s) leads to only small changes in the solution. Here's two well posed learning problems presented.

Ad-sensitive customers

The retailer store chain Mall-Wart want to install screens that present various offers to their customers. A screen should be placed next to a cash register, so that an offer is presented to the customer when the cashier is in the process of scanning their products. The offer should be chosen such that the customer are most likely to accept it.

This can be posed as a machine learning problem:

- Task T: decide which offer to present to the customer

- Performance measure P: percent of accepted offers

- Training experience E: The product set, the chosen offer, and whether the offer was accepted or not, are stored in a database on every sale.

Self-driving shopping cart robot

Mall-Wart want their customers to be followed around by a self-driving shopping cart. This can also be posed as a machine learning problem:

- Task T: navigating through the store by following a customer around using vision sensors

- Performance measure P: average distance traveled before the customer correct the robot

- Training experience E: A visual record of the store together with the navigation patterns of a human doing the task.

1.2

What is inductive bias (for a learning method)? Why is it so important in machine learning? The candidate elimination algorithm for learning in version spaces and learning of decision trees with ID3 are two different learning methods. What can you say about the inductive bias for each of them?

Let X be the set of instances of a concept learning problem, and let H be the set of hypotheses a learner may consider regarding learning the identity of the target concept c. **Inductive bias** for a learning method L are the assumptions the learner make that limit H, so that the cardinality $|H|$ of the hypotheses set are smaller than the cardinality of the power set of X, i.e $|H| < |\mathcal{P}(X)|$.

Inductive bias is an important concept in machine learning because a biased learner can make inductive leaps to classify unseen examples, x_i. For

example, the inductive bias of the CANDIDATE-ELIMINATION algorithm is the assumption that $c \in H$. The bias exhibited by the ID3 algorithm is less precise, nevertheless one can approximate its bias as saying ID3 prefers shorter (simpler) trees (hypotheses) over complex trees. In addition, ID3 prefers trees where information is clustered close to the root.

We see that in both cases the learner has a bias for selection one consistent hypothesis over another.

2   Problem Two

2.1

- x_1: number of black pieces next to each other
- x_2: number of white pieces next to each other
- x_3: number of two available squares next to the black pieces
- x_3: number of two available squares next to the white pieces
- x_3: number of available squares next to and in line with the black pieces
- x_3: number of available squares next to and in line with the white pieces

2.2

As a 4x4 matrix. In python:

```
B = 3*[3*[0]]
```

2.3

2.4

2.5

2.6

2.7

3   Coding

```
def foo(bar):
    print 'Eat my ' + bar
```

```python
1  def foo(bar='Chocolate'):
2      rage = 'Did you eat my ' + bar + '?!'
3      print rage
```

Remember to

```python
import gravity
```

4   Citing

This is a citation[1]. And you can use the citation verbally as well, according to Kiss et al. [1].

5   Mathemagic

An inline equation is written $x^3 = 8$. You can also use this

$$x^5 + 4x^2 = 1337x \tag{1}$$

to write a 1337 equation.

References

[1] Gabriel Kiss, Erik Steen, Jon Petter Helgesen Å sen, and Hans G Torp. GPU volume rendering in 3D echocardiography: Real-time pre-processing and ray-casting. In *2010 IEEE International Ultrasonics Symposium*, pages 193–196. IEEE, October 2010. ISBN 978-1-4577-0382-9. doi: 10.1109/ULTSYM.2010.5935485. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5935485`.